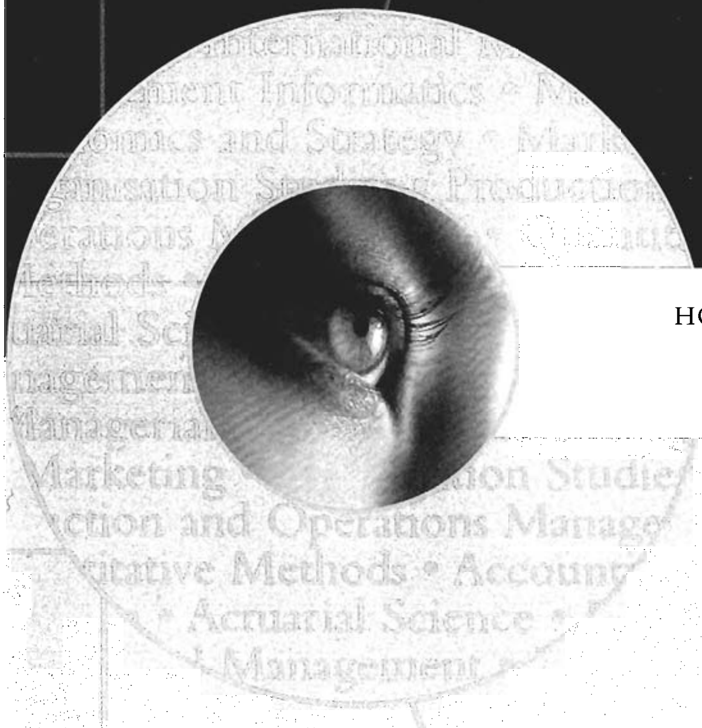
DEPARTMENT OF APPLIED ECONOMICS



RESEARCH REPORT

HOW THE CO-INTEGRATION ANALYSIS CAN HELP IN MORTALITY FORECASTING

GRZEGORZ DARKIEWICZ & TOM HOEDEMAKERS

OR 0406

How the Co-Integration Analysis Can Help in Mortality Forecasting

Grzegorz Darkiewicz[§] Tom Hoedemakers[†]

Actuarial Science Research Group
Catholic University of Leuven
Naamsestraat 69, 3000 Leuven
Belgium
tel: 0032 16 32 67 61
fax: 0032 16 32 67 32
[§]Grzegorz.Darkiewicz@econ.kuleuven.ac.be
[†] Tom.Hoedemakers@econ.kuleuven.ac.be

How the Co-Integration Analysis Can Help in Mortality Forecasting

Abstract

The method of mortality forecasting proposed in 1992 by Lee and Carter describes a time series of age-specific log-mortality rates as a sum of an independent of time age-specific component and a bilinear term in which one of the component is a time-varying parameter reflecting general change in mortality and the second one is an age-specific factor. Such a rigid model structure implies that on average the mortality improvements for different age groups should be proportional, regardless the calendar period.

In this paper we investigate whether the mortality data for England and Wales follow this property or not. We perform the analysis by applying the concept of the Engle and Granger co-integration to the time series of log-mortality rates. We investigate the goodness of fit of the predictions to the historical data. We find that a lack of co-integration indeed can cause some problems in performance of the model. In the last section we propose several opportunities to omit the pitfalls.

Keywords: the Lee-Carter model, time series analysis

1 Introduction and motivation

During the 20^{th} century the life time expectancy increased dramatically - for example for England and Wales in 1900 the life expectancy at birth was 48.15 years for females and 44.23 for males, while in 1995 - 79.46 for females and 74.25 for males (the source: Human Mortality Database [11]). Usually we consider the mortality improvements as something positive and optimistic - we live statistically longer than our ancestors. On the other hand when we think about the assumptions of modern social security systems, such dramatic changes in the mortality may be also seen as one of the major threats to them. Thus they pose a great challenge for actuaries, especially those planning public retirement systems and private life annuities business. In fact all the components of social security systems are affected by mortality trends. Therefore nowadays reasonable mortality forecasting techniques are of paramount importance for the society.

In the 20^{th} century global mortality has declined at relatively constant rate. However significant heterogeneity was observed in a number of deaths by age, a cause of deaths and a calendar year. When one chooses an appropriate model for forecasting future trends, one must foresee whether the model would reflect this heterogeneity. One must also rise a more fundamental question: is using historical data theoretically sound at all? It is well-known that the mortality in the previous centuries declined much slower than in the 20^{th} century. Can we thus assume that present trends will stand on for the next decades? One has also to determine whether arbitrarily small mortality can be reached in the model or rather some biological barriers should be imposed. All these questions undermine the sense of forecasting mortality in a very long time perspective. However for average time horizons such forecasts are necessary, so *nolens-volens* one has to choose the most suitable forecasting method.

It has been empirically tested that the rate of improvement is age- and gender- specific, and thus most of modern methodologies concern the mortality rates separately for both genders and different ages. There are several approaches to develop suitable models. Some parametric methods can be easily obtained in the framework of Generalized Linear Models. It has been argued that the number of deaths when the central exposed-to-risk is given may be assumed to follow Poisson distribution (see [3]) and the promising estimates may be obtained by fitting the Poisson regression (see [19], [17] and [20]). An interesting alternative was proposed in 1992 by Lee and Carter ([13]) who developed a method combining parametric approach with time series analysis. Recently the Lee-Carter model has been widely discussed in actuarial literature. Some essential improvements were introduced by Brouhns *et al.* ([4]) who estimated parameters by Poisson log-bilinear regression and Renshaw and Haberman ([18]) who described the model in the GLM terms.

In this contribution we evaluate performance of the Lee-Carter model from another perspective. In the first part of the analysis we examine whether age-specific log-mortality rates for England and Wales for years 1901-1995 are pairwise co-integrated. In the second part we make forecasts for years 1971-1995 based on the same data restricted to years 1901-1970, and compare them to the historical data to test the efficiency of the model.

In Section 2 we briefly describe the assumptions of the Lee-Carter methodology and the estimation methods used to obtain the forecasts. Section 3 explains the relationship between the assumptions of the Lee-Carter model and the concept of co-integration. The data sources are described in Section 4. The results of the co-integration analysis are presented in Section 5. In Section 6 we compare obtained estimates and forecasts to the historical data. Next we make some suggestions about possible ways of improving the classical Lee-Carter methodology in Section 7. Finally, Section 8 briefly summarizes the paper.

2 The Lee-Carter methodology with some modifications

The model proposed in [13] (see also [12]) is a very powerful and elegant approach to mortality projections. It specifies log-linear form for the force of mortality $\mu_x(t)$. More precisely, in the model the following relation is assumed:

$$\ln \hat{\mu}_x(t) = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt}, \tag{1}$$

where $\hat{\mu}_x(t)$ denotes the estimated mortality rate for people at age x in calendar year t, ε_{xt} - an error term, in classical approach assumed to be homoskedastic (the estimation methods considered more recently in actuarial literature, e.g. [4] allow to release this assumption), α_x describes the shape of the age profile (can be computed for example by averaging over time), β_x - the pattern of deviations from the age profile, and κ_t is an age-independent process describing time-deviations of mortality. The mortality rates are estimated here as a ratio of an actual number of deaths D_{xt} to a central exposed-to-risk E_{xt}.

One can easily check that the structure is invariant under either of the parameter transformations:

$$\Big\{\alpha_x, \beta_x, \kappa_x\Big\} \mapsto \Big\{\alpha_x, \frac{\beta_x}{c}, c\kappa_x\Big\},$$

$$\big\{\alpha_x, \beta_x, \kappa_x\big\} \mapsto \big\{\alpha_x - c\beta_x, \beta_x, \kappa_x + c\big\}.$$

Usually for uniqueness of the model specification following constraints are imposed:

$$\sum_t \kappa_t = 0 \text{ and } \sum_x \beta_x = 0.$$

In classical settings parameters α_x, β_x and κ_t were estimated by minimizing the sum of squares:

$$\sum_{x,t} \big(\ln \hat{\mu}_x(t) - \alpha_x - \beta_x \kappa_t\big)^2.$$

The estimation problem cannot be solved by a simple regression model because of the presence of a bilinear term. The minimization of the sum consists of taking $\hat{\alpha}_x$ as a raw average of $\ln \hat{\mu}_x(t)$'s and then getting $\hat{\beta}_x$ and $\hat{\kappa}_t$ from the first term of singular value decomposition (SVD) of the matrix $\big[\ln \hat{\mu}_x(t) - \hat{\alpha}_x(t)\big]_{xt}$. Next the values κ_t are re-estimated (taken $\hat{\alpha_x}$ and $\hat{\beta}_x$ as given) so that the following identity holds:

$$\sum_x D_{xt} = \sum_x E_{xt} \exp\big(\hat{\alpha}_x + \hat{\beta}_x \hat{\hat{\kappa}}_t\big).$$

This means that after re-estimation the resulting death rates applied to actual exposures-to-risk will produce total number of deaths actually observed each year.

The estimated time-dependent parameter $\hat{\hat{\kappa}}_t$ can be seen as a stochastic process. Then the forecasts can be obtained by modeling $\hat{\hat{\kappa}}_t$ as an ARIMA(p,q,s) process, using standard Box and Jenkins methodology (identification-estimation-diagnosis) (see [2]). Denoting the resulting projections beyond the data time horizon T as $\hat{\hat{\kappa}}_{T+s}$, the forecasted mortality rates will be expressed by the formula:

$$\hat{\mu}_x(T+s) = \hat{\mu}_x(T) \cdot \exp\Big(\beta_x\big(\hat{\hat{\kappa}}_{T+s} - \hat{\hat{\kappa}}_T\big)\Big).$$

However, as pointed out in [1], the classical methodology of estimating parameters imposes too restrictive conditions on the error structure in equation (1). For this reason in our numerical illustration we will adopt the Poisson log-bilinear regression developed in [4].

The method assumes that the number of deaths of people at age x in year t is Poisson-distributed (according to [3] this assumption is plausible), namely

$$D_{xt} \sim \text{Poisson}\big(E_{xt}\mu_x(t)\big), \text{ where } \ln \mu_x(t) = \alpha_x + \beta_x \kappa_t. \tag{2}$$

The parameters α_x, β_x and κ_t are estimated by maximizing the Poisson log-likelihood function, which takes the following form :

$$L(\alpha, \beta, \kappa) = \sum_{x,t} \big(D_{xt}(\alpha_x + \beta_x \kappa_t) - E_{xt} \exp(\alpha_x + \beta_x \kappa_t)\big) + \text{constant}.$$

Because of the presence of the bilinear term $\beta_x \kappa_t$, in our estimations one has to use numerical procedures. Following [4], we use an iterative method proposed in [10], which is based on the following general scheme:

$$\hat{\theta}^{(\nu+1)} = \hat{\theta}^{(\nu)} - \frac{\frac{\partial L}{\partial \theta}\left(\hat{\theta}^{(\nu)}\right)}{\frac{\partial^2 L}{\partial \theta^2}\left(\hat{\theta}^{(\nu)}\right)}.$$

This leads to the following explicit algorithm:

$$\hat{\alpha}_x^{(\nu+1)} = \hat{\alpha}_x^{(\nu)} - \frac{\sum_t (D_{xt} - \hat{D}_{xt}^{(\nu)})}{-\sum_t \hat{D}_{xt}^{(\nu)}}, \quad ,\hat{\beta}_x^{(\nu+1)} = \hat{\beta}_x^{(\nu)}, \quad ,\hat{\beta}_x^{(\nu+1)} = \hat{\beta}_x^{(\nu)},$$

$$\hat{\kappa}_t^{(\nu+2)} = \hat{\kappa}_t^{(\nu+1)} - \frac{\sum_x (D_{xt} - \hat{D}_{xt}^{(\nu+1)})\hat{\beta}_x^{(\nu+1)}}{-\sum_t \hat{D}_{xt}^{(\nu+1)}(\hat{\beta}_x^{(\nu+1)})^2}, \quad ,\hat{\alpha}_x^{(\nu+2)} = \hat{\alpha}_x^{(\nu+1)}, \quad ,\hat{\beta}_x^{(\nu+2)} = \hat{\beta}_x^{(\nu+1)},$$

$$\hat{\beta}_x^{(\nu+3)} = \hat{\beta}_x^{(\nu+2)} - \frac{\sum_t (D_{xt} - \hat{D}_{xt}^{(\nu+2)})\hat{\kappa}_t^{(\nu+2)}}{-\sum_t \hat{D}_{xt}^{(\nu)}(\hat{\kappa}_t^{(\nu+2)})^2}, \quad ,\hat{\alpha}_x^{(\nu+3)} = \hat{\alpha}_x^{(\nu+2)}, \quad ,\hat{\kappa}_t^{(\nu+3)} = \hat{\kappa}_t^{(\nu+2)},$$

where $\hat{D}_{xt}^{(\nu)} = E_{xt} \exp\left(\hat{\alpha}_x^{(\nu)} + \hat{\beta}_x^{(\nu)}\hat{\kappa}_t^{(\nu)}\right)$. As starting values we have taken $\hat{\alpha}_x^{(0)} = 0$, $\hat{\beta}_x^{(0)} = 1$, $\hat{\kappa}_t^{(0)} = 0$ and we have stopped the iteration when the increase in log-likelihood function after all three steps was sufficiently smaller than 10^{-4}.

3 The concept of co-integration and its relations with the Lee-Carter model

Suppose that one has have two time series variables X_t and Y_t, which can be decomposed as follows:

$$X_t = a(t) + u_t, \tag{3}$$

$$Y_t = b(t) + v_t, \tag{4}$$

where processes $a(\cdot)$ and $b(\cdot)$ represent non-stationary time trends and u_t, v_t - the irregular stationary components. One says that variables X_t and Y_t are co-integrated if there exist non-zero values β_1 and β_2 such that the linear combination $\beta_1 X_t + \beta_2 Y_t$ is stationary, which means that the term $\beta_1 a(t) + \beta_2 b(t)$ has to vanish.

The co-integration analysis is usually performed in economic sciences to determine whether there exist some unique relationships between economical variables resulting in a long-term equilibrium.

In this contribution we deploy the method of co-integration analysis developed by Engle and Granger ([8]). Their testing methodology proceeds in two steps. In the first step it has to be verified whether the variables under consideration are indeed non-stationary. The non-stationarity can be stated by means of so-called unit root tests. Usually it is required that the variables have exactly one unit root (i.e. the first differences are stationary)

The most popular method of testing the existence of unit roots is the Augmented Dickey-Fuller test (ADF) (see [6], [7]). One tests the hypothesis of unit root against the alternative hypothesis that the series is autoregressive of order $k+1$ ($AR(k+1)$). In the ADF test the following equation is deployed:

$$X_t - X_{t-1} = bX_{t-1} + \sum_{j=1}^{k} c_j(X_{t-j} - X_{t-j-1}) + \varepsilon_t, \tag{5}$$

where ε_t is assumed to be a white noise process and k denotes the number of lagged first difference terms. In standard applications there are two modifications of the test; the first one including the constant term:

$$X_t - X_{t-1} = c + bX_{t-1} + \sum_{j=1}^{k} c_j(X_{t-j} - X_{t-j-1}) + \varepsilon_t \tag{6}$$

and the second additionally including the trend variable:

$$X_t - X_{t-1} = c + at + bX_{t-1} + \sum_{j=1}^{k} c_j(X_{t-j} - X_{t-j-1}) + \varepsilon_t. \tag{7}$$

The test relies on rejecting the null hypothesis of the unit root ($H_0 : b = 0$) in favor of stationarity. To test this hypothesis, a negative and significant (non-normally distributed) t-ratio for b has to be computed and then compared to critical values reported in [6] or more recently in [14]. If the hypothesis of the unit root cannot be rejected, the test is repeated for first differences to check the existence of multiple unit roots (one has to determine whether the order of integration of the tested variables is equal exactly to 1).

The Phillips-Perron (PP) test (see [16]) is an alternative approach to test existence of unit roots. While the ADF test corrects for higher order serial correlation by adding lagged difference terms on the right-hand side, the PP test makes the correction to the t-statistic of the b coefficient for one of the AR(1) regressions of the form (5), (6) or (7) (i.e. when k is equal to 0). More precisely, the following equation is employed:

$$X_t - X_{t-1} = \alpha + \beta X_{t-1} + \varepsilon_t$$

(with possible modifications when there is no intercept term and when we additionally consider a trend variable). The PP test is robust to heteroskedasticity and autocorrelation of unknown form of $\{\varepsilon_t\}$. In our application we deploy tests provided by EViews, which are based on the Newey-West correction (see [15]). The asymptotic distribution of the PP t-statistic is the same as the ADF t-statistic, thus its value is again compared to the critical values reported in [6] or [14].

After the existence of unit roots has been stated for variables X_t and Y_t, one has to verify whether co-integrating constants β_1 and β_2 exist (it can be assumed that β_1=1). This is done by performing two symmetric OLS estimations:

$$Y_t = a_2 + b_2 X_t + u_{2t} \tag{8}$$

$$X_t = a_1 + b_1 Y_t + u_{1t}, \tag{9}$$

and testing the stationarity of u_{1t} and u_{2t} by the Augmented Dickey-Fuller test (in this case however the values of the t-statistic are compared to critical values reported in [9]). If the unit root hypothesis is rejected for u_{1t} then one can take $\beta_1 = 1$ and $\beta_2 = -b_1$, and as a consequence $\beta_1 X_t + \beta_2 Y_t = a_1 + u_{1t}$ is stationary. The analogous reasoning may be carried out for the equation (9).

Now let us return to the Lee-Carter model. We will consider log-mortality rates as a set of time series variables indexed by age $\{\ln \hat{\mu}_x(t)\}_x$ (note that we perform whole analysis for both genders separately). According to the equation (1), the Lee-Carter model assumes the long-term relationship between log-mortality rates and a common co-integrating variable κ_t. In some sense this representation is similar to (3) and (4). Indeed, consider two ages x_1 and x_2 and let $X_1 = \log \mu_{x_1}(t)$ and $X_2 = \log \mu_{x_2}(t)$. Then

$$X_1 = \beta x_1 \kappa_t + \alpha_{x_1} + \varepsilon_{x_1 t}$$

and

$$Y_1 = \beta x_2 \kappa_t + \alpha_{x_2} + \varepsilon_{x_2 t}.$$

Note that in the original methodology of Lee and Carter the assumptions on error terms ε_{xt} were very close to stationarity (homoskedasticity of variance and mean reversion). Despite in more recent works these assumptions are not so strict and allow even for some systematic patterns (the approach of [4]), the most important property of a stationary process, i.e. mean reversion, should be satisfied. Moreover a high number of parameters in the model imposes that the variability of error terms should be relatively small. Thus if the model is specified correctly it may be assumed that a possible co-integration relationship will not be affected by an error structure.

For these reasons it is not the best practice to check the stationarity of the residuals directly. Their shape heavily depends on the employed estimation methodology. Moreover the systematic patterns which may appear in the time series variables describing error structure may result in rejecting the stationarity hypothesis even if their real impact on long term relationships between log-mortality rates is negligible..

In exchange it seems to be a much better idea to test whether log-mortality rates for different ages are co-integrated. Indeed, consider two ages x_1 and x_2. Then the long-term relationship will be given by the formula:

$$\ln \hat{\mu}_{x_1}(t) = \alpha_{x_1} - \frac{\beta_{x_1}}{\beta_{x_2}}\alpha_{x_2} + \frac{\beta_{x_1}}{\beta_{x_2}} \ln \hat{\mu}_{x_2}(t) + \varepsilon_{x_1 t} - \frac{\beta_{x_1}}{\beta_{x_2}}\varepsilon_{x_2 t}. \tag{10}$$

If the error terms $\varepsilon_{x_1 t}$ and $\varepsilon_{x_2 t}$ are stationary than the co-integration follows immediately. If not - it is still very likely that co-integrating constants between the series of log-mortality rates can be found independently on the error structure given by a specific estimation model. Moreover, we expect that it should be much easier to find co-integrating relations for all possible pairs than to find one co-integrating process κ_t for all log-mortality rates simultaneously (mathematically these properties are equivalent, but from the statistical point of view pairwise tests are much weaker).

Summarizing, the assumptions of the Lee-Carter model and the Engle-Granger co-integration, despite not mathematically equivalent, have many points of tangency. In fact the logic of the Lee-Carter model is based on the observation that time changes of mortality for log-mortality rates for different ages have always the same (up to an error term) proportions, regardless the calendar period. Despite the Engle and Granger co-integration analysis is formulated in a bit different language, we are convinced that it provides a very useful tool to make the diagnostic checks of validity of the Lee-Carter model. In this paper we illustrate our findings by applying the Lee-Carter method and co-integration tests to the 20^{th} century mortality data for England and Wales.

4 The description and sources of the data

The analysis is performed on the basis of population estimates and death counts for England and Wales in the period 1901-1995. More exactly, we use death counts D_{xt} for years 1901-1995 and all ages between 0 and 110+ years, as well as the estimates of exposure-to-risk E_{xt} and mortality rates $\hat{\mu}_x(t)$. All data are provided separately for both genders.

The original data come from the following sources:

1. Population estimates:

    - Office for National Statistics (1998). "Twentieth Century Mortality in England and Wales" (CD-ROM). Newport, South Wales: Office for National Statistics.

- Office for National Statistics. Population estimates unit. Unpublished data.

2. Death counts:

   - Philipov, D. "Construction of the England and Wales population and mortality surfaces, 1841-1999". Unpublished manuscript.
   - Title of tables: "Deaths at Different Ages". Registrar's General Annual Report, 1901-1910.
   - General Register Office (1911-1920). "Annual Report of the Registrar General". London: Her Majesty's Stationery Office.
   - General Register Office (1921-1973). "Registrar General's Statistical Review of England and Wales". London: Her Majesty's Stationery Office.
   - Office of Population Censuses and Surveys (1974-1995). "Mortality Statistics" (Series DH1). London: Her Majesty's Stationery Office.

The data were downloaded through the Human Mortality Database on 14 April 2003. In our analysis we used also estimates of exposure-to-risk and death rates obtained by HMD.

5 The co-integration analysis for log-mortality rates

In this section we investigate whether age-specific log-mortality rates for England and Wales are pairwise co-integrated. We perform the tests for all combinations of five different ages: 5, 25, 40, 60 and 75 years, separately for males and females. We proceed with Engle and Granger's procedure in two steps, as described in Section 3.

5.1 Testing for unit roots

More careful analysis of the data indicates that log-mortality rates for England and Wales reveal significant variations for years of both world wars ($1914-1918$, $1939-1944$) and epidemics (Spanish flu in 1918). Also in 1929 an unexpected increase in mortality was noted. Thus the assumption of heteroskedasticity and serial independence of error terms in the formula (5) is very difficult to satisfy even for a very large number of lagged differences on the right-hand side of the equation. For these reasons we use for our purpose Phillips-Perron test instead of Augmented Dickey-Fuller, for which the conditions for error terms are less rigid.

In Table 1 there are numerical results of the test presented.

Table 1: Values of the t-statistic for log-mortality rates $m_x(t) = \ln(\hat{\mu}_x(t))$

	Males		Females	
x	$m_x(t)$	$\nabla m_x(t)$	$m_x(t)$	$\nabla m_x(t)$
5	-0.237162	-15.69261	0.007477	-12.39218
25	-0.440493	-13.15585	-1.368653	-7.811529
40	-0.835487	-16.98648	-0.894117	-11.30519
60	-0.239665	-16.35535	0.528027	-13.35696
75	-0.136543	-18.49173	-0.844876	-17.01046

We compare the results to the critical values from Table 2. The hypothesis of existence of unit roots cannot be rejected for neither of tested time series variables. For first differences there is a clear indication of stationarity (the hypothesis of unit root is easily rejected for all variables). We conclude that all variables are integrated of order 1, and thus the assumptions necessary to proceed with estimating the co-integrating equations (8) and (9) are satisfied.

5.2 The tests for co-integration

We deploy the procedure of Engle and Granger described in Section 3. After OLS-estimation of (8) and (9), we test the stationarity of the residual series using the equation (5). Preliminarily we choose the number of lagged differences k which minimizes the Akaike Information Criterion (AIC). Next we check by usual Q-Statistic if the residuals are not serially correlated. If the hypothesis of white noise is not rejected we use in the model the number k, otherwise - we aim to choose the smallest $k' > k$ such that residuals from the equation (5) are not serially correlated. Because of a very big sensitivity of the results to the choice of the model, in ambiguous cases we also report the results for the model with increased number of lagged differences.

The values of t-statistic are contained in Table 3. We use the following labels in the table: $x\mathrm{S}y$ means the co-integration test for the sex S (where S means "M" for males and "F" for females) of log-mortality rates for ages x and y. In the third column we report results when log-mortality

Table 2: The critical values reported in [14]

	Confidence level		
	90%	95%	99%
$m_x(t)$	-2.5829	-2.8922	-3.5007
$\nabla m_x(t)$	-2.5831	-2.8925	-3.5015

Table 3: The results of the co-integration tests

	Lags	Eq. (8)	Eq. (9)
$5M25$	1	-2.609717	-2.635994
$5M40$	2	-2.752887	-2.943339
	3[1]	-2.593201	-2.848422
$5M60$	1	-2.751095	-2.780564
$5M75$	2	-2.754339	-2.550786
$25M40$	1	-2.407018	-2.495128
$25M60$	1	-2.220739	-2.200334
$25M75$	1	-2.743430	-2.566435
$40M60$	1	-3.034872	-2.976308
	2 [1]	-2.710061	-2.637798
$40M75$	9	-1.696533	-0.925303
$60M75$	9	-2.414889	-1.867568
$5F25$	1	-3.152333	-3.587565
$5F40$	0	-3.136495	-3.284172
	1[1]	-2.360561	-2.579673
$5F60$	1	-1.511879	-1.248186
$5F75$	2	-2.340452	-2.000002
$25F40$	0	-3.030436	-2.927567
	1 [1]	-3.393295	-3.149962
$25F60$	0	-2.043662	-1.565188
$25F75$	2	-2.130491	-1.004689
$40F60$	3	-0.506666	0.212523
$40F75$	4	-2.040707	-1.520733
$60F75$	2	-2.599192	-2.373252

for age x is the independent variable in the equation (8) and log-mortality for age y dependent, and in the fourth column the opposite case. The values of t-statistic are compared to the Engle and Yoo critical values reported in Table 4.

5.3 Conclusions

The analysis of results contained in Table 3 reveals that for most of the tested pairs log-mortality rates are not co-integrated. The results strongly support co-integration for only two pairs: females aged 5 with females aged 25 (for the equation (9) the test rejects the hypothesis of

[1] An explaining test

Table 4: The critical values reported in [9]

	Confidence level		
	90%	95%	99%
No lags	-3.03	-3.37	-4.07
Lags	-2.91	-3.17	-3.73

a unit root at the confidence level 5%, while for the equation(8) at 10%) and females aged 25 with females aged 40 (the hypothesis is rejected only for the equation (8) at the confidence level 10%. The explaining tests with one additional lagged difference in the equation 5 reject the hypothesis once again - for (8) at 5% and for (8) at 10%). The results for females aged 5 with females aged 40 also support the co-integration - the hypothesis of a unit root is rejected at 10% both for (8) and (9) (explaining tests did not allow to reject the hypothesis). Note that these results are consistent with the theoretical property of transitivity of the co-integration relation.

From all remaining pairs only the tests for males aged 40 with males aged 60 allow for rejecting the hypothesis of unit root at the level 10% (however explaining tests did not allow for rejecting the hypothesis). For remaining sixteen combinations neither of 32 tests allowed for rejecting the null hypothesis. Although the p-values of the tests usually seem to be relatively small, the results make the assumption of co-integration of log-mortality rates for all ages doubtful, at least in the case the tested data set.

The results of the tests suggest that the Lee-Carter methodology, is not fully applicable to the 20^{th} century mortality data for England and Wales. In the next section we compare the predictions obtained from the Lee-Carter forecasts to the historical data. It is possible to notice that indeed the proportions of mortality improvements between different ages do vary with time, what is linked to the lack of co-integration. In Section 7 we discuss how it is possible to modify the Lee-Carter methodology to make the mortality forecasts more reliable.

6 The forecasts obtained by the model

Apart from the co-integration analysis, we make also the forecasts to look at the results of the model. The forecasts are derived for the period 1971-1995 on the basis of the mortality data for England and Wales for years 1901-1970. Then the estimates for the period 1901-1995 are compared graphically to the historical data. We employ the methodology of [4] described in Section 2.

Table 5: Estimated parameters of the model (11)

	Males		Females	
	Coef.	St.er.	Coef.	St.er.
C	-0.011035	0.001597	-0.010099	0.001761
λ	-0.501670	0.101670	-0.369395	0.114339

The raw estimates of α_x, β_x and κ_t are inserted in the Appendix. However obtained estimates of κ_t are not easy to model as an ARIMA process because of an excessive variability of mortality in the periods of wars (1914-1918, 1939-1944) and epidemics (1918 and probably 1929). Therefore we used the smoothed process $\widetilde{\kappa}_t$ obtained from the following formula:

$$\widetilde{\kappa}_t = \begin{cases} \frac{1}{6}\big((1919-t)\kappa_{1913} + (t-1913)\kappa_{1919}\big) & \text{for } t = 1914, \ldots, 1918 \\ \frac{1}{2}\big(\kappa_{1928} + \kappa_{1930}\big) & \text{for } t = 1929 \\ \frac{1}{7}\big((1945-t)\kappa_{1938} + (t-1938)\kappa_{1945}\big) & \text{for } t = 1939, \ldots, 1944 \\ \kappa_t & \text{otherwise} \end{cases}$$

After these adjustments the Box and Jenkins methodology (identification - estimation - diagnosis) was employed to generate an appropriate ARIMA time series model for mortality index $\widetilde{\kappa}_t$. Both indices for males and females were modelled as ARIMA(1,1,0) process, i.e.:

$$\widetilde{\kappa}_t - \widetilde{\kappa}_{t-1} = C + \lambda(\widetilde{\kappa}_{t-1} - \widetilde{\kappa}_{t-2}) + \varepsilon_t, \tag{11}$$

where ε_t forms a white noise process. In Table 5 we insert the estimated parameters.

We depict the results on two sets of graphs. In Figure 1 we depict the historical evolution of the mortality rates for chosen ages, both for the historical data and for the Lee-Carter estimates and forecasts. In Figure 2 the global age-specific log-mortality rates are depicted for chosen calendar years.

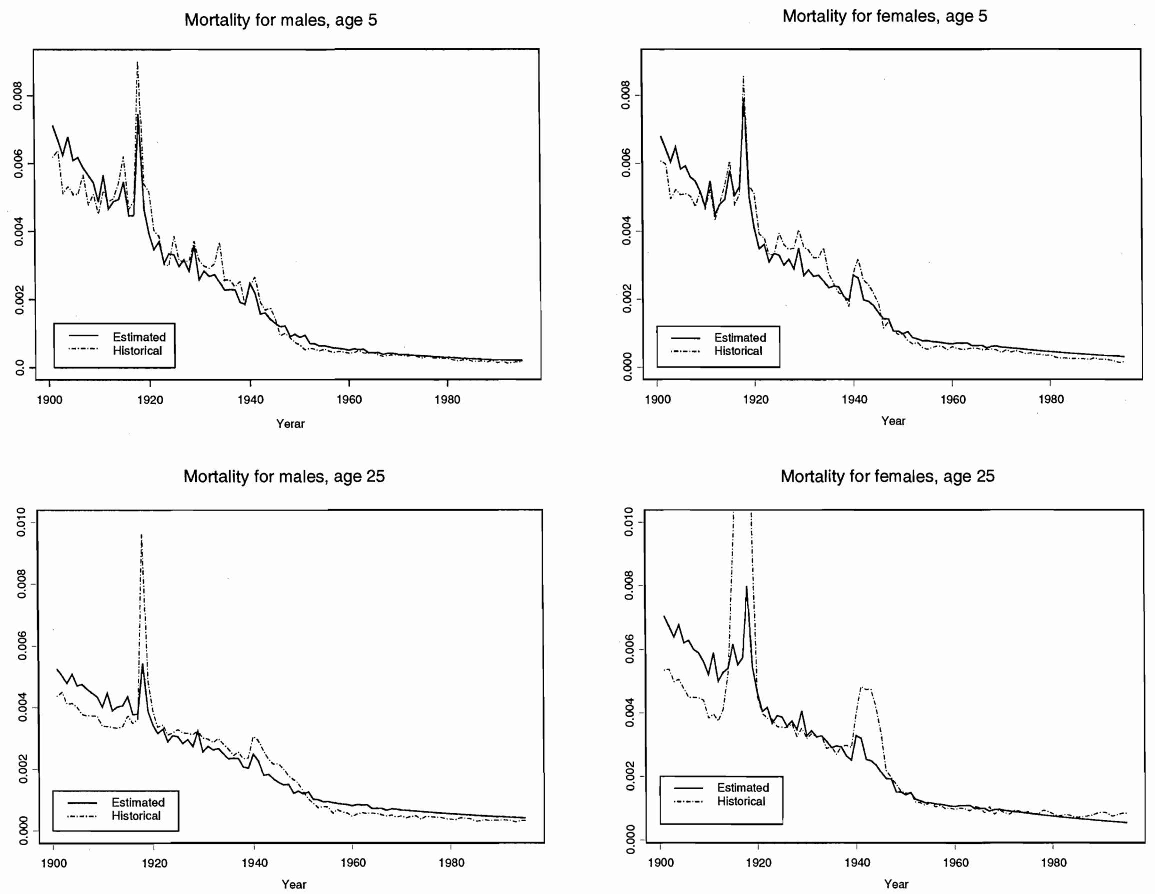
At a first view the fit of the Lee-Carter estimates to the historical data seems to be reasonably good. However the lack of co-integration leads to several inconsistencies. For the year 1951 for example the model seems to overestimate mortality for males aged between 5 and 30. This tendency is kept for the following years, but the predictions for 1995 do not reveal it any more. However then the mortality for elderly males is overestimated significantly. This may suggest that from 1970 the pace of improvement for represented by the parameters β_x should be decreased for ages 20-30 while should be increased for elderly ages. For females this phenomenon is illustrated even more clearly. For most of the years the fit for females is even better than for males. However on the last graph of Figure 2 (i.e. for year 1995) the fit for females is very bad - the mortality is significantly underestimated for years 20-40 and overestimated for elderly women. Those phenomena result from the fact that the assumption of constant β_x is not always plausible and thus the long-term relationship (10) does not hold.

We want to stress that in short time despite these problems the forecasts still may perform reasonably good. Moreover in some applications (for example reserving in life-annuity business) overestimation of the mortality for some ages may be compensated by underestimation for others. However the lack of co-integration of the log-mortality series suggests that the model is not enough flexible and that it cannot be used in very long perspective (the example for females that already 25-year age-specific forecasts turn out to be very inadequate).

7 How to omit pitfalls?

The Lee-Carter model can be made more efficient in several ways.

One of possible reasons for which log-mortality rates for different ages may not be pairwise co-integrated is too long time perspective. Indeed, in the classical Lee-Carter model the same



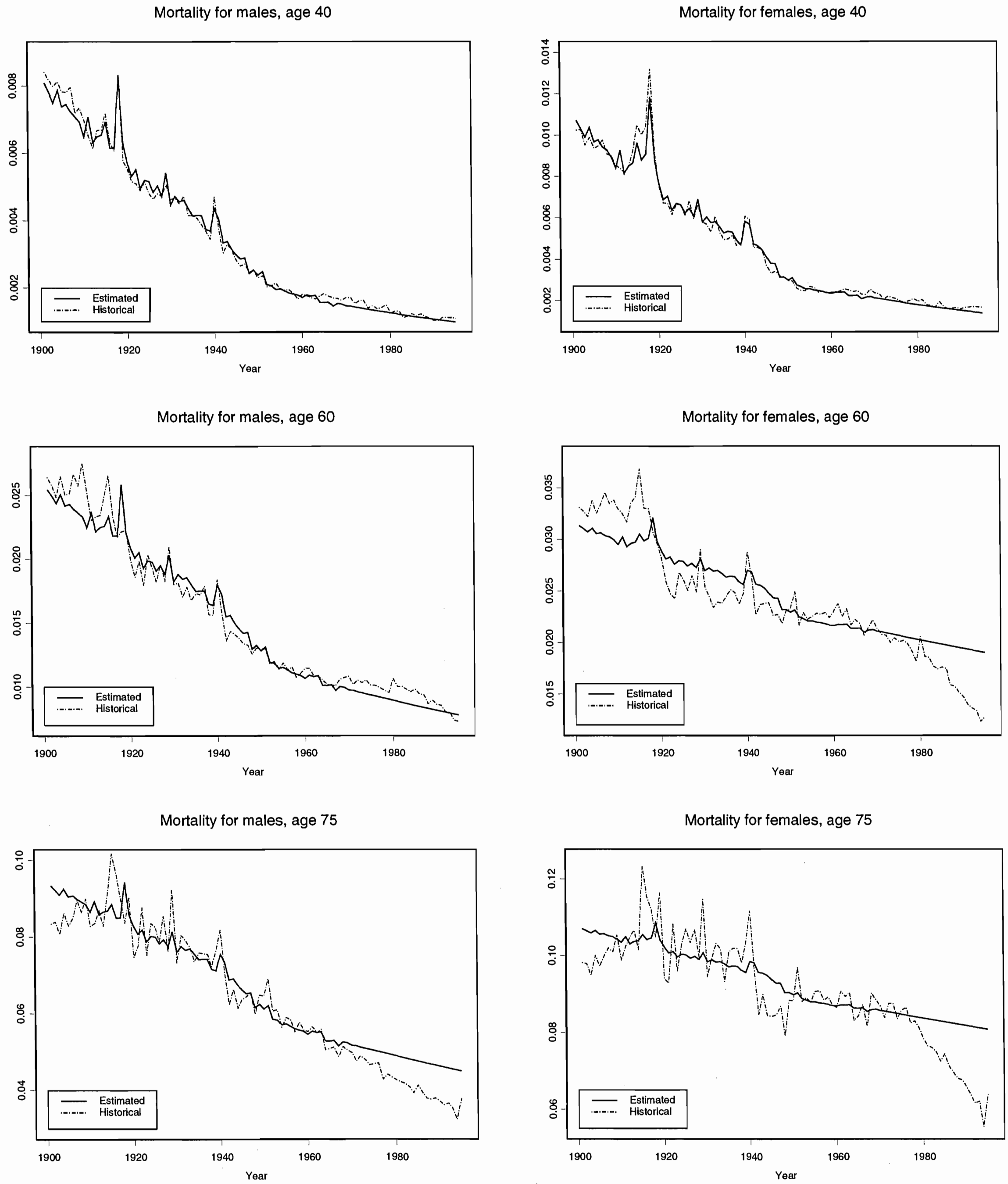


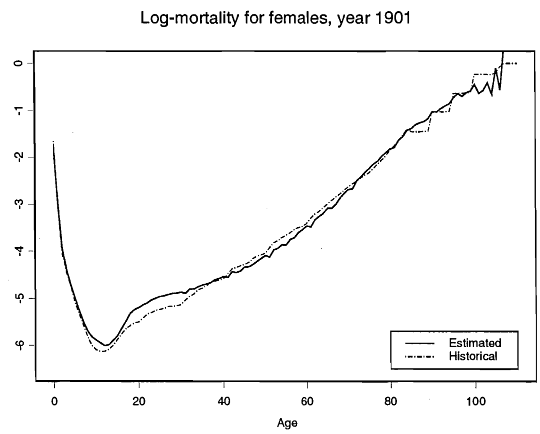
Figure 1: Changes of Mortality for England and Wales over Time for Chosen Ages

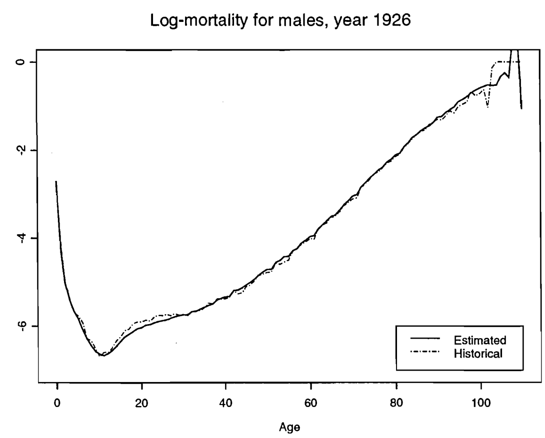
weight is put to the observations at the beginning as at the end of the period. It does not always reflect the reality - it is well known that for example mortality trends in the thirties were influenced mostly by improvements in mortality caused by infectious diseases, from which infants and young people benefited relatively more than elderly people. Thus probably the time period taken for the analysis is too long. It is also possible to use similar approach to this of

Renshaw and Haberman ([18]). In their generalized linear modelling based regression approach to mortality forecasting they propose to add a time break-point for greater structural flexibility. Translating their idea into the classical Lee-Carter model settings, the addition of the break-point means that we choose a time point t_0 and estimate two sets of parameters: β_x for $t < t_0$ and β'_x for $t \geq t_0$. The motivation is to put greater emphasis on more recent trends. Obviously such an approach will produce better fit to the historical data, but on the other hand the modification substantially increases number of parameters involved.

Also disaggregation of the data may lead to a substantial improvement of the results. There are two possible ways of disaggregation. For the first one all calculations are performed for every group separately, in particular death rates are modelled separately. The disaggregation with respect to gender is of this type. Geographical disaggregation is another example. A division of the population into smaller age-groups is also possible (for example for calculating the present value of life annuity the analysis can be restricted to elderly people).

In some cases it is possible to disaggregate only death counts, keeping the central exposed-to-risk E_{xt} at global level. The death rates in disaggregation with respect to an underlying cause of





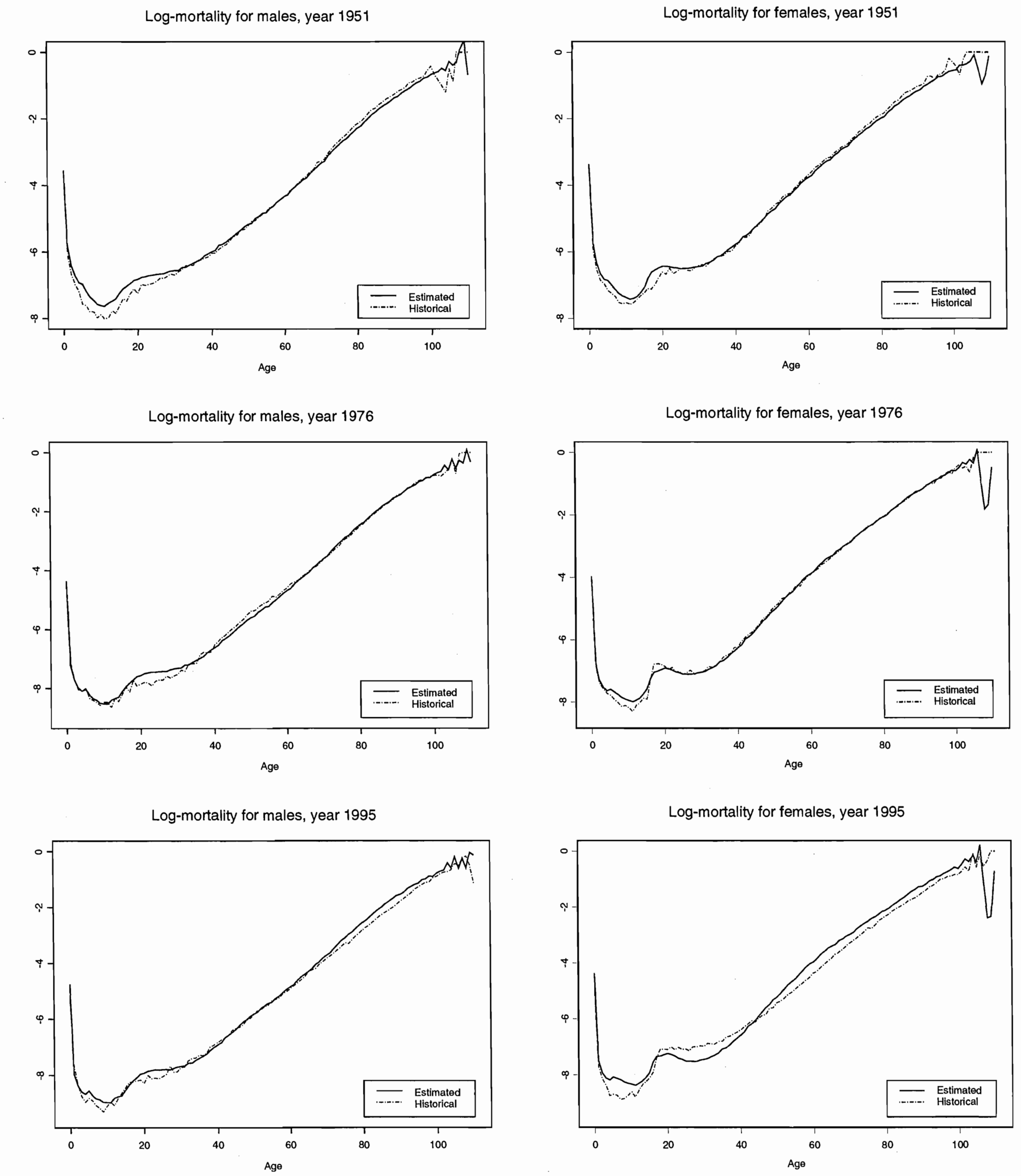


Figure 2: Log-Mortality Rates for England and Wales for Chosen Calendar Years

death may be performed this way. The main advantage of this approach is that the death rates modeled separately can be added together to get global estimates of death rates.

$$\hat{\mu}_x^{(i)}(t) = \frac{D_{xt}^{(i)}}{E_{xt}} \text{ for } i = 1, \dots n,$$

and thus

$$\hat{\mu}_x(t) = \frac{\sum_{i=1}^{n} D_{xt}^{(i)}}{E_{xt}} = \sum_{i=1}^{n} \hat{\mu}_x^{(i)}(t).$$

However one has to remember to take possible dependencies between the numbers of deaths between different subgroups into account.

The common feature of both types of disaggregation is estimation of separate sets of parameters $\alpha_x^{(i)}$, $\alpha_x^{(i)}$ and $\kappa_t^{(i)}$ which may catch some specific differences in mortality for analyzed subgroups. Especially disaggregation with respect to the cause of death seems to be a very promising improvement in the model. The main advantage of this approach is that it really explains the reasons of the lack of co-integration, unlike other methods which aim only to improve the fit by some manipulation with the data or parameters. Indeed, while in the first half of the 20^{th} century the mortality improvements followed from rapid decrease in the mortality caused by infectious diseases, in the second half of the century improvements resulted mainly from decrease in number of deaths caused by cardiovascular diseases (see e.g. [5]). These facts could explain why in the first years of the century infants benefited the most from mortality improvements, while recently especially elderly people did.

Disaggregation with respect to the cause of death has also some disadvantages. The results of such forecasts cannot be reliable in a very long time perspective because usually old causes of death are replaced by new ones, for example recently AIDS became the one which should be analyzed separately. Also proper data are often unavailable.

However we consider examining disaggregation with respect to the cause of death as an interesting topic for future research which can increase our knowledge about behavior of mortality rates in time.

8 Summary

In the paper we have made an attempt to evaluate the Lee-Carter model of forecasting future mortality. In the analysis we have used the concept of the Engle and Granger co-integration and have applied it pairwise to the log-mortality rates.

We have performed the analysis for 20^{th} century data for England and Wales. The tests that we have used did not allow for stating perfect pairwise co-integration between age-specific log-mortality rates what undermined the reliability of the Lee-Carter model for this particular data set. The comparison of the Lee-Carter forecasts with the historical data seems to confirm conclusions derived from the co-integration analysis.

In practical applications we suggest making co-integration tests before deploying the Lee-Carter model as a method of diagnostic checking. The lack of co-integration may be a kind of warning signal that the obtained predictions may not be reliable. It may be also an indication that more weight should be put on more recent observations in the model or that a more disaggregated analysis is required, for example with respect to a cause of death.

Acknowledgements

The authors acknowledge financial support by the Onderzoeksfonds K.U.Leuven (GOA/02: Actuariële, Financiële en Statistische Aspecten van Afhankelijkheden in Verzekerings- en Financiële Portefeuilles).

References

[1] Alho, J.M. "Discussion of Lee (2000)." *North American Actuarial Journal* 4 (2000): 91-93.

[2] Box, G.E.P., Jenkins, G.M. *Time Series Analysis for Forecasting and Control.* San Francisco: Holden-Day, 1976.

[3] Brillinger, D.R. "The Natural Variability of Vital Rates and Associated Statistics." *Biometrics* 42 (1986): 693-734.

[4] Brouhns, N., Denuit, M., Vermunt, J.K. "A Poisson Log-Bilinear Regression Approach to the Construction of Projected Life Tables." *Insurance: Mathematics and Economics* 31 (2002): 373-393.

[5] Cutler, D.M., Meara, E. "Changes in the Age Distribution of Mortality over the 20^{th} Century." *National Bureau of Economic Research Working Paper* 8556 (2001).

[6] Dickey, D.A., Fuller, W.A. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74 (1979): 427-431.

[7] Dickey, D.A., Fuller, W.A. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica* 49 (1981): 1057-1072.

[8] Engle, R.F., Granger, C.W.J. "Co-Integration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55 (1987): 251-276.

[9] Engle, R.F., Yoo, B.S. "Forecasting and Testing in Cointegrated Systems." *Journal of Econometrics* 35 (1987): 143-159.

[10] Goodman, L.A. "Simple Models for the Analysis of Association in Cross-Classification Having Ordered Categories." *Journal of the American Statistical Association* 74 (1979): 537-552.

[11] *Human Mortality Database.* University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de.

[12] Lee, R. "The Lee-Carter Method of Forecasting Mortality, with Various Extensions and Applications." *North American Actuarial Journal* 4 (2000): 80-93.

[13] Lee, R.D., Carter, L.R. "Modelling and Forecasting U.S. Mortality." *Journal of the American Statistical Association* 87, 419 (1992): 659-671.

[14] MacKinnon, J. "Critical Values for Co-Integration Tests." *UCSD Economic Discussion Papers*, University of California at San Diego, Department of Economics, 1990: 90-94.

[15] Newey, W.K., West, K.D. "A Simple Positive Semi-Definite, Hederoskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica* 55 (1987): 703-708.

[16] Phillips, P.C.B., Perron, P. "Testing for a Unit Root in Time Series Regression." *Biometrica* 75 (1988): 335-346.

[17] Renshaw, A., Haberman, S. "Dual Modelling and Select Mortality." *Insurance: Mathematics and Economics* 19, 2 (1997): 105-126.

[18] Renshaw, A., Haberman, S. "Lee-Carter Mortality Forecasting: a Parallel Generalized Linear Approach for England and Wales Mortality Projections." *Applied Statistics* 52, 1 (2003): 119-137.

[19] Renshaw, A., Haberman, S., Hatzoupoulos, P. "The Modelling of Recent Mortality Trends in United Kingdom Male Assured Lives", *British Actuarial Journal* 2, I (1996): 185-205.

[20] Sithole, T.Z., Haberman, S., Verrall, R.J. "An Investigation into Parametric Models for Mortality Projections, with Applications to Immediate Annuitants' and Life Office Pensioners' Data", *Insurance: Mathematics and Economics* 27 (2000): 185-312.

Appendix

1. Estimations of κ_t

Year	Males	Females	Year	Males	Females
1901	-0.97430559	-0.958104082	1931	-1.216216948	-1.205971396
1902	-0.99027241	-0.974269994	1932	-1.231579354	-1.225761779
1903	-1.008666032	-0.992290267	1933	-1.226009201	-1.221947415
1904	-0.986613478	-0.972095934	1934	-1.249588039	-1.24118375
1905	-1.015198594	-1.002461809	1935	-1.27425508	-1.265146559
1906	-1.011420245	-0.998083218	1936	-1.272134314	-1.258777394
1907	-1.024521078	-1.014063799	1937	-1.273252529	-1.262292297
1908	-1.034699659	-1.020292117	1938	-1.319566742	-1.296961205
1909	-1.044487163	-1.03672619	1939	-1.326620825	-1.31504895
1910	-1.073830793	-1.062839727	1940	-1.251860885	-1.222294395
1911	-1.034785807	-1.020388652	1941	-1.285961335	-1.231432388
1912	-1.085226528	-1.077676314	1942	-1.370988783	-1.312498501
1913	-1.073419504	-1.058270761	1943	-1.365532173	-1.318329955
1914	-1.069837085	-1.049796109	1944	-1.397726122	-1.337062245
1915	-1.043832253	-1.004458296	1945	-1.421440492	-1.372975934
1916	-1.096844574	-1.043184417	1946	-1.441280236	-1.406624158
1917	-1.096122439	-1.02936574	1947	-1.437027614	-1.407601818
1918	-0.961929332	-0.914803598	1948	-1.514874179	-1.491054728
1919	-1.086127366	-1.045447573	1949	-1.495239337	-1.491003813
1920	-1.130387126	-1.102782603	1950	-1.52149409	-1.511645528
1921	-1.162554216	-1.149898735	1951	-1.503266271	-1.495100076
1922	-1.146006203	-1.139435872	1952	-1.580296797	-1.546229624
1923	-1.195304455	-1.183502215	1953	-1.585247293	-1.560233067
1924	-1.172252865	-1.161512647	1954	-1.611651573	-1.578428701
1925	-1.175760205	-1.166099545	1955	-1.607191768	-1.578142882
1926	-1.204205221	-1.192292801	1956	-1.621742746	-1.589252898
1927	-1.187098462	-1.176541902	1957	-1.63714992	-1.592647397
1928	-1.216307455	-1.202971375	1958	-1.645503413	-1.601395485
1929	-1.154164596	-1.147982019	1959	-1.652902364	-1.610483914
1930	-1.241019442	-1.223324843	1960	-1.670147532	-1.614248975

Year	Males	Females	Year	Males	Females
1961	-1.651045988	-1.605536763	1966	-1.708026196	-1.63308312
1962	-1.660026179	-1.60752804	1967	-1.743908519	-1.667533585
1963	-1.655077903	-1.603523958	1968	-1.718527936	-1.651148081
1964	-1.71234977	-1.636153364	1969	-1.723902053	-1.646427648
1965	-1.712736555	-1.6362825	1970	-1.74107159	-1.660457184

2. Estimations of α_x

Age	Males	Females	Age	Males	Females
0	0.756053556	0.989470086	25	-2.618123183	-2.187871525
1	2.065584527	1.927293013	26	-2.569170566	-2.156596711
2	0.642585944	0.442440266	27	-2.474258357	-2.090826926
3	-0.056939177	-0.404445804	28	-2.489758885	-2.054592584
4	-0.60176891	-0.992395348	29	-2.486965391	-2.093039163
5	-1.231736069	-1.649877077	30	-2.463555998	-2.067280226
6	-1.597217368	-2.166509278	31	-2.51628299	-2.156029884
7	-1.976874715	-2.622173856	32	-2.466719417	-2.043714119
8	-2.345146713	-2.928718298	33	-2.468521947	-2.096655593
9	-2.535604251	-3.129531053	34	-2.492381124	-2.075411218
10	-2.682436986	-3.234558488	35	-2.491742029	-2.152612356
11	-2.769015215	-3.316128665	36	-2.494318601	-2.143782159
12	-2.912388993	-3.519230046	37	-2.529897612	-2.207388642
13	-2.81073017	-3.560274224	38	-2.50561661	-2.214076203
14	-2.621847699	-3.504850055	39	-2.578247045	-2.269947499
15	-2.660962679	-3.498853361	40	-2.640107858	-2.323920979
16	-2.706419281	-3.559600207	41	-2.784827008	-2.493601998
17	-2.756150577	-3.484473791	42	-2.71538618	-2.459624784
18	-2.803776125	-3.13222094	43	-2.841033286	-2.585584107
19	-2.84754822	-3.027223007	44	-2.871609613	-2.645821125
20	-2.839806971	-2.983211785	45	-2.876352502	-2.641792413
21	-2.794461194	-2.820436423	46	-2.944958424	-2.795475708
22	-2.802444332	-2.679274844	47	-2.943273963	-2.84999956
23	-2.747762913	-2.455011166	48	-2.850948271	-2.830100846
24	-2.649795248	-2.337182194	49	-2.84985682	-2.886257435

Age	Males	Females	Age	Males	Females
50	-2.830479078	-2.859598421	81	-1.447956531	-1.535117502
51	-2.920377532	-3.077658117	82	-1.259945996	-1.33519572
52	-2.722915918	-2.934552033	83	-1.206159532	-1.2610996
53	-2.75708249	-3.025012285	84	-1.060233473	-1.094390792
54	-2.680836986	-2.995977392	85	-1.053235918	-1.094107666
55	-2.741115685	-3.116786454	86	-1.033222028	-1.047895064
56	-2.612001896	-3.023807562	87	-1.000163626	-1.075140726
57	-2.641512709	-3.09629136	88	-1.011784611	-1.093620018
58	-2.518886503	-3.005258398	89	-1.003650919	-1.069586163
59	-2.523936701	-3.003759608	90	-0.730636124	-0.768945901
60	-2.450118222	-2.922193913	91	-0.837120518	-0.869461423
61	-2.537013474	-3.076665257	92	-0.771152519	-0.836690787
62	-2.38017569	-2.920782536	93	-0.782204824	-0.79188261
63	-2.345716215	-2.90798614	94	-0.748975568	-0.797089968
64	-2.251069339	-2.885118341	95	-0.497347545	-0.557844779
65	-2.143988007	-2.74004098	96	-0.565194394	-0.471521647
66	-2.253222003	-2.809382133	97	-0.446303037	-0.660900162
67	-2.23038108	-2.770711407	98	-0.381883865	-0.575932491
68	-2.062470375	-2.563753424	99	-0.185188857	-0.605379471
69	-2.0372435	-2.462474372	100	-0.259950057	-0.263674126
70	-1.926539384	-2.367066795	101	-0.274953832	-0.825585094
71	-1.969225896	-2.412187376	102	-0.247872349	-0.888274508
72	-1.781608491	-2.187799819	103	-0.716481988	-0.450301995
73	-1.752335052	-2.126817077	104	-0.380329639	-1.200233179
74	-1.696292135	-2.014982473	105	-0.550012318	0.206897013
75	-1.618168838	-1.92954453	106	0.343152439	-1.410040803
76	-1.574112216	-1.842844623	107	-0.522729747	2.189181678
77	-1.539421461	-1.783325386	108	2.645506697	5.20672437
78	-1.456607765	-1.655215243	109	1.792922708	6.356951022
79	-1.470775409	-1.61323437	110+	-2.62658354	2.448484316
80	-1.388068186	-1.517785324			

3. Estimations of β_x

Age	Males	Females	Age	Males	Females
0	2.870401684	2.913072841	31	2.68455006	2.854036591
1	5.216765725	5.10800461	32	2.658489669	2.877434088
2	4.688123047	4.541651873	33	2.653474505	2.818507253
3	4.437988277	4.179973631	34	2.590594417	2.79087059
4	4.200743893	3.901039783	35	2.555851624	2.673408706
5	3.811970384	3.487184153	36	2.517727121	2.655986224
6	3.712842329	3.233028808	37	2.459854478	2.5623581
7	3.569920374	3.012425102	38	2.405943439	2.48596007
8	3.396962869	2.890248793	39	2.31542088	2.401039231
9	3.347082747	2.805224655	40	2.233565783	2.309657515
10	3.276249224	2.766283401	41	2.106197709	2.153380692
11	3.233779502	2.746850032	42	2.051312695	2.064417918
12	3.077109445	2.596203364	43	1.940485735	1.947351374
13	3.10236347	2.530357151	44	1.876636567	1.849843222
14	3.185258706	2.492292501	45	1.814249666	1.766626314
15	3.05473092	2.396584646	46	1.710007788	1.601369025
16	2.932146671	2.161410026	47	1.647546991	1.491875825
17	2.838435766	2.088514245	48	1.650957922	1.441129216
18	2.749856164	2.276233621	49	1.588422152	1.319890898
19	2.661044028	2.307506848	50	1.556665896	1.279723598
20	2.647419376	2.307531403	51	1.479994965	1.091179088
21	2.637732283	2.415029379	52	1.513632009	1.08052535
22	2.619631506	2.520079422	53	1.449363053	0.951172494
23	2.635053194	2.687613338	54	1.438194781	0.90260531
24	2.690590012	2.771477785	55	1.385469485	0.781555858
25	2.699827377	2.887974365	56	1.3840364	0.753707096
26	2.721700886	2.907316256	57	1.31848553	0.640219817
27	2.778241328	2.947292344	58	1.319820894	0.620952584
28	2.738040757	2.962989023	59	1.260400918	0.558342881
29	2.718249298	2.915918786	60	1.252542649	0.564673399
30	2.719940721	2.918294442	61	1.164103669	0.417177637

Age	Males	Females	Age	Males	Females
62	1.161829235	0.424717737	87	0.412504043	0.202418015
63	1.11751987	0.368704643	88	0.36035656	0.152774421
64	1.119194834	0.314638088	89	0.30664939	0.111180424
65	1.125208861	0.358667404	90	0.428339164	0.271792408
66	1.016097113	0.287974555	91	0.329506584	0.172074105
67	0.942115805	0.229080195	92	0.304044911	0.130458584
68	0.981052357	0.310798857	93	0.247350948	0.120423047
69	0.926452422	0.312500355	94	0.227790933	0.066523154
70	0.932025454	0.332176287	95	0.33623781	0.177727565
71	0.863947718	0.269928349	96	0.237809179	0.176447266
72	0.876687875	0.318730092	97	0.28519342	0.043463286
73	0.821342702	0.300974732	98	0.264711414	0.048917908
74	0.784225142	0.311821767	99	0.37607998	-0.016978934
75	0.772566106	0.31810442	100	0.273892899	0.194977313
76	0.728184621	0.309767294	101	0.231370662	-0.194900536
77	0.706442223	0.308746539	102	0.224784737	-0.326617114
78	0.679328301	0.327691158	103	-0.153818917	-0.038898005
79	0.611130919	0.297276697	104	0.116671104	-0.566556224
80	0.600127216	0.31826416	105	-0.17729093	0.321766145
81	0.518872707	0.267347987	106	0.491622899	-0.88664423
82	0.552485686	0.313549317	107	-0.144562406	1.804259818
83	0.521029993	0.309261661	108	1.685894202	4.120525857
84	0.539979253	0.346761892	109	0.953475578	4.715852723
85	0.494637644	0.307375602	110+	-1.295952254	1.716894008
86	0.440343638	0.271616394			